

Una metodología para el análisis geoespacial de mortalidad asociada a contaminación industrial en México

Prefacio

El mapeo de las enfermedades ha sido una parte fundamental de las luchas ambientales a nivel mundial. Existen casos destacados en Latinoamérica como el Barrio de Ituzaingó en Córdoba, Argentina (La Nueva Mañana, 2017), donde la comunidad mapeó la zona para investigar potenciales asociaciones entre las fumigaciones aéreas de agroquímicos con la incidencia de cáncer y malformaciones congénitas; otro ejemplo es el estudio epidemiológico liderado por Damián Verzeñassi (Livingstone, 2016) quien realizó un censo de 80,000 habitantes en Rosario, Argentina, con el cual logró establecer una asociación entre la contaminación agroindustrial y los problemas de salud en la región. Ambas experiencias demuestran que las enfermedades en una comunidad específica no son un mero conjunto de individuos enfermos, sino una determinación social, “un proceso metabólico con la naturaleza”.¹

Según Breilh (Breilh 1989), la epidemiología se ha convertido en tan sólo un “aditamento secundario” de la clínica, reduciendo la habilidad explicativa de aquella a un conjunto de reglas “objetivas” que derivan de un método positivista. Aunque la estadística y los métodos matemático-analíticos son necesarios e importantes, es imperante recordar que el saber de las comunidades y su forma particular de aprehender su territorio contienen validez explicativa de los fenómenos, incluyendo las patologías de la región.

El análisis epidemiológico del proceso de “salud-enfermedad” es una “unidad de la realidad social de la naturaleza y la realidad natural del hombre” (Breilh 1989). En otras palabras, no es únicamente la dieta o el consumo de drogas lo que se asocia, respectivamente, con la obesidad o las adicciones, sino que las enfermedades y el proceso de salud responden necesariamente a la

¹ Véanse los escritos filosóficos de Marx (1982) y el tomo primero de El Capital (1975). Para Marx, la relación del ser humano con la naturaleza es un proceso metabólico mediado por el trabajo y por la forma histórica en la que ese trabajo se apropia de lo natural para satisfacer necesidades.

relación del metabolismo social con la naturaleza y la producción social del hábitat natural en el que vive el ser humano. Por lo tanto, no es fortuito que algunas comunidades en México hayan señalado que el deterioro significativo en la salud de sus habitantes comenzara en el momento en que la gran industria empezara a transformar su hábitat, imponiéndose en sus territorios, articulándose entre ellos con carreteras, o desechando residuos en ríos que antes servían no sólo para beber e irrigar campos de cultivo, sino como articulador de culturas, expresión de tradiciones e inspiración estética. La crisis socio-ambiental actual ha propiciado importantes luchas sociales que exigen que esta se reconozca como un problema prioritario de salud pública, así como la necesidad de intervenciones urgentes para revertir y prevenir el deterioro de su entorno y su salud.

Existen pocas investigaciones científicas que analicen el problema del deterioro sistemático del medio ambiente y la salud de la población desde la perspectiva del modo particular de producción del capitalismo. Esta metodología pretende llenar este vacío del conocimiento, generando información relevante para las comunidades en lucha por la devastación ambiental. Mi principal inspiración para desarrollarla fue la *Coordinadora por un Atoyac con vida*, una organización en el estado de Tlaxcala, México, que ha luchado incansablemente por recuperar los ríos que alguna vez fueron fuente de vida y patrimonio de sus comunidades. La *Coordinadora* entendió que la contaminación industrial del río no sólo afectaba la tierra, el suelo y sus cultivos, sino la salud, la cultura, y las tradiciones de su comunidad. En pocas palabras, trastocaba la viabilidad misma de las comunidades de la región.

Las enfermedades neoplásicas y crónico-degenerativas en Tlaxcala son altamente prevalentes en la cuenca Atoyac-Zahuapan (Comisión Nacional de Derechos Humanos, 2017). Los métodos aquí vertidos han permitido mapear la mortalidad por insuficiencia renal, el cáncer, y la diabetes, así como los abortos espontáneos en esta y otras regiones críticamente afectadas del país. La utilización del análisis geoespacial y los modelos de elevación permiten conocer el riesgo

de fallecer por estas enfermedades según la proximidad de las localidades a una gran industria. Cabe destacar que las comunidades de la cuenca habían identificado este riesgo por décadas sin que las autoridades locales prestaran atención, y mucho menos que lo establecieran como una prioridad de salud pública.

Por último, se enfatiza que la Ciencia en México debe tener un papel preponderante en la re-estructuración de las políticas de salud y la recuperación de las comunidades afectadas. La Ciencia debe partir del principio revolucionario del conocimiento, de la ampliación de la conciencia humana para todos y del desarrollo de las capacidades colectivas para resolver la crisis civilizatoria de la que el capitalismo ya no puede salir. Cualquier actividad científica que no parta de principios éticos, comunitarios, deliberativos, críticos, autocríticos y el respeto a los saberes y conocimientos de otras culturas, pueblos y tradiciones es una producción poco útil que corre el riesgo de propiciar más injusticias de las que pretende resolver. Basándose en estos principios, este proyecto surgió de una larga lucha de las comunidades en México y de un intenso y nutrido dialogo entre éstas e investigadores mexicanos, para desarrollar una nueva metodología que utiliza diversas disciplinas científicas como la epidemiología, la estadística, y la informática. Se espera que sirva como parteaguas a muchas otras comunidades en su lucha contra la rápida y progresiva degradación ambiental.

Samuel Rosado Zaidi

Facultad de Economía, UNAM

21 de agosto de 2018

1. Fuentes de información

La información fue obtenida de las bases de datos del Instituto Nacional de Estadística y Geografía, México (INEGI), en específico de las “Defunciones generales” del año 2002 al 2016 que están disponibles en el sitio web del Instituto (INEGI 2016). Dichas bases de datos registran las defunciones a nivel nacional por persona de manera anónima e incluye, entre otras variables, la edad del individuo fallecido, la causa básica de muerte consignada en el certificado de defunción, así como la entidad, municipio y localidad de residencia. Las causas de defunción se enlistan como códigos de la Clasificación Internacional de Enfermedades CIE-10 (WHO, 2016) las cuales se desglosan en un diccionario de datos del INEGI. A fin de establecer una asociación entre las causas de defunción y las características de la zona de residencia del individuo fallecido se utilizó la entidad, municipio y localidad de residencia y no la entidad de registro y ocurrencia de la defunción. Los datos de población se obtuvieron del censo 2010, también disponible en la página web del INEGI (INEGI 2012). El censo incluye a la población correspondiente a 192,247 localidades del país, con datos desagregados por sexo y agrupados por edad.

2. Unificación de bases de datos y asignación de claves geográficas

En la primera fase del procesamiento de los datos se importaron los datos del censo y las causas de defunción al software R versión 3.5. Para presentar los datos de manera geoespacial, se combinaron los códigos del INEGI para la entidad federativa, el municipio y la localidad para generar una clave geográfica para cada individuo, la cual se utilizó como variable en las bases de datos de defunciones. Posteriormente, se procedió a unir las bases de datos de defunciones con el censo 2010. Esto permite relacionar el número de defunciones por una enfermedad específica con una localidad determinada, tanto cartográfica como demográficamente. Los mapas se generaron con el software QGIS versión 2.18.

Para analizar las bases de datos fue necesario procesarlas con un listado de códigos. Es importante destacar que la unificación de las dos bases de datos, por sí sola, no produce resultados suficientes

ni estadísticamente significativos.

Para producir la Clave Geográfica se concatenaron las columnas "ENT_RESID", "MUN_RESID", "LOC_RESID"

```
> DF$CVEGEO = paste0(DF$ent_resid, DF$mun_resid, DF$loc_resid)
```

Produciendo la siguiente clave:

```
> head(DF$CVGEO.C.9)
[1] "020040001" "300399999" "050350001" "090070001" "140867777"
```

Para la estandarización de datos, cada clave geográfica debe contar con nueve dígitos: la clave del estado debe contener dos dígitos; para los estados con un dígito se debe agregar un cero al principio. La clave del municipio debe contener tres dígitos y las localidades, cuatro dígitos. De no ser así, los espacios deben ser rellenados con ceros. Este punto es sumamente importante para poder unir la información de las distintas bases de datos. Se puede emplear el comando `str_pad` del paquete `stringr` de R (Wickham 2018) para la verificación de los nueve dígitos.

Dado que las causas de defunción específicas (columna: `causa_def`) y las causas de defunción por lista especial de tabulación (columna: `lista1`) están expresadas como códigos, fue necesario unir las columnas de códigos con el catálogo de datos de la INEGI que arroja la descripción de la enfermedad. Esto permitió visualizar el nombre de las causas de mortalidad y realizar subconjuntos más confiables.

```
> head(DF$causa_def)
[1] P220 P072 P220 P280 P072 P072
3936 Levels: A010 A020 A021 A028 A029 A039 A044 A047 A048 A049 A058
A059 ... Y899
```

Se unió con la siguiente función:

```
merge(DF, causas, by.x=c("causa_def"), by.y=c("CVE"))
```

3. *Selección de las defunciones por cáncer*

Se utilizó el paquete reshape2 (Wickham, 2007) para generar conteos de las causas de defunción para cada grupo de edad por localidad determinada. Ya que la lista completa de neoplasias en la base de datos de INEGI es muy amplia, se agruparon en grandes grupos por sitio como se señala en la Tabla 1. Se generó un subconjunto de los datos con el comando subset, filtrando con los códigos en la columna “CAUSA_DEF” por la letra C.

Tabla 1. Grupos de neoplasias que se incluyeron en el análisis de mortalidad por cáncer en México

| |
|--|
| Tumores malignos del labio, de la cavidad bucal y de la faringe |
| Tumor maligno del esófago |
| Tumor maligno del estómago |
| Tumor maligno del colon, del recto y del ano |
| Tumor maligno del hígado y de las vías biliares intrahepáticas |
| Tumor maligno del páncreas |
| Tumor maligno de la laringe |
| Tumor maligno de la tráquea, de los bronquios y del pulmón |
| Melanoma maligno de la piel |
| Tumor maligno de la mama |
| Tumor maligno del cuello del útero |
| Tumor maligno de otras partes y de las no especificadas del útero |
| Tumor maligno del ovario |
| Tumor maligno de la próstata |
| Tumor maligno de la vejiga urinaria |
| Tumor maligno de las meninges, del encéfalo y de otras partes del sistema nervioso central |
| Linfoma no Hodgkin |
| Mieloma múltiple y tumores malignos de células plasmáticas |
| Leucemia |
| Resto de tumores malignos |
| Resto de tumores |

4. Cálculo de las Razones Estandarizadas de Mortalidad

Dado que hay localidades pequeñas en el país con un número muy reducido de defunciones para realizar los análisis estadísticos se agruparon los datos de mortalidad por los siguientes lustros: 2002-2006, 2007-2011, 2012-2016.

Para cada localidad se calculó una razón estandarizada de mortalidad (REM) con la siguiente

$$\text{fórmula: } REM_i = \frac{O_{i(l)}}{E_{i(l)}}$$

En la que $O_{i(l)}$ son las muertes observadas (en este caso, mortalidad específica por cáncer) para los siguientes grupos de edad: 0 a 2 años, 2 a 5 años, 6 a 11 años, 12 a 14 años, 15 a 17 años, 18 a 24

años, 25 a 29 años y de 60 en adelante durante los lustros (2002-2006, 2007-2011, 2012-2016) en la localidad i . $E_{i(l)}$ son las muertes esperadas para los grupos de edad y lustros correspondientes de acuerdo a las tasas crudas de mortalidad de la población mexicana. Las edades no fueron seleccionadas acorde a los estándares internacionales de división de edades debido a que el censo de población no permite esa extrapolación a nivel localidad. Para ajustar por estas diferencias, se realizó una regresión con el método empírico de Bayes acorde a recomendaciones de otros investigadores (Piñeros, et al, 2010; Razaeeian, et al, 2007) para suavizar las cifras producto del azar, muy por encima de lo esperado o en poblaciones pequeñas (Gómez-Rubio, et al, 2005). Se realizaron dos regresiones: en la primera, se dejaron las muertes esperadas por debajo de 1 sin alterar dado que esto altera la regresión empírica de Bayes en el caso de poblaciones muy pequeñas (por ejemplo, localidades con una muerte por lustro pueden salir muy por encima de lo esperado si esta población es muy pequeña). Posteriormente, se realizó una segunda regresión para las localidades en las cuales la mortalidad esperada por cáncer fuese menor a 1; para aquellas localidades donde la mortalidad esperada fuese menor a uno y la mortalidad observada mayor a 1, la mortalidad esperada se establecía en 1, sin importar que fuesen menores de 0.5; para las localidades donde la mortalidad observada fuese menor a 11 y la mortalidad esperada menor a 1, la mortalidad esperada se establecía en 1 para las que fuesen 0.5 en adelante y a 0 todas las que fuesen menor a 0.5². Todas las localidades sin datos se asumieron como una razón estandarizada de mortalidad de 0 para poder utilizar todas las localidades con coordenadas en el proceso que se describe más adelante.

Es importante destacar que se realizaron ajustes de población básicos, tomando en consideración las proyecciones globales de crecimiento poblacional, para ajustar la mortalidad esperada para cada localidad. Se usó la tasa de crecimiento (tomando como base la población de 2010) para cada municipio y se aplicó a cada localidad. Debido a que los microdatos del censo no están disponibles

² Tomando en consideración los lineamientos establecidos por Washington State Department of Health (2012b), en los casos donde hay menos de 12 defunciones el estadístico en cuestión es inestable y no debe ser presentado. No obstante, se presentan todos los estadísticos y se etiquetan aquellos inestables.

públicamente, no es posible realizar un ajuste poblacional estadístico. Ya que no existen datos desglosados para todas las localidades del país como aquellos disponibles en el censo, utilizaron los datos de mortalidad de 2009 a 2013 sin un ajuste de población.

5. Uso de las estadísticas de distancia G_i y el análisis del vecino más cercano (*nearest neighbor analysis*)

Dada la heterogeneidad que implica el análisis de localidades tan distintas y dispersas en el país, se realizó un análisis de asociación espacial que permite identificar agrupamientos geográficos definidos como “pockets” por Getis y Ord (Getis y Ord 1992). La metodología propuesta por estos autores construye un estadístico $G_i(d)$ que calcula la relación entre la distancia de los “pockets” y una variable x , como puede ser el número de homicidios o la tasa de mortalidad. El estadístico es una proporción de las sumas de todos los valores dentro de una distancia de x :

$$G_i(d) = \frac{\sum_{j=1}^n w_{ij}(d) x_j}{\sum_{j=1}^n x_j}$$

Para este caso, se evalúan las razones estandarizadas de mortalidad (REM) que están dentro de la distancia d de i . Por lo tanto, si hay valores altos de x en tal distancia d_i , entonces el estadístico $G_i(d)$ es alto. Del mismo modo, esta metodología construye un estadístico Z_i que mide la variación de los valores $G_i(d)$ con respecto a la media de los valores de x . En pocas palabras, si el estadístico Z_i es muy alto implica que los valores de x son más altos que la media de x , y que están dentro de la distancia d de i . Por el contrario, valores pequeños de Z_i implican que los valores de x son menores a la media (*ibid*).

Como señalan los autores, el estadístico suaviza los errores asociados con la densidad de puntos agrupados en la “vecindad”. Si el valor esperado de una localidad i es el mismo para todas las x , se neutraliza el patrón “pocket”. El valor del estadístico $G_i(d)$ puede ser alto por la proximidad de las localidades, pero sólo en donde la variable x se aleje sistemáticamente de la media es que se

pueden identificar patrones locales de asociación.

Utilizando los métodos de Bivand (2017), se identificaron los puntos circundantes unos a otros con $k=1, 2$ y 4 , es decir, por distancia de su vecino más próximo. Además, se utilizó la metodología de vecinos por distancia y se evaluaron los vecinos a $1, 2$ y 5 kilómetros. Al haber diversas localidades sin registro de defunción por cáncer y para poder mostrar la mayoría de las localidades en el mapa se asignaron valores de 0 a las localidades que no registraron defunciones por cáncer. Después de extraer las coordenadas y los nombres de las filas se utilizó el siguiente comando para agrupar a los “vecinos” (Bivand et al, 2015):

```
>VECINOS <- knn2nb(knearneigh(coords, k = 1), row.names = Ids)
>VECINOS <- include.self(VECINOS)
>VECINOS2 <- dnearneigh(CENSO, d1=0, d2=2, row.names=Ids)
>VECINOS2 <- include.self(VECINOS2)
```

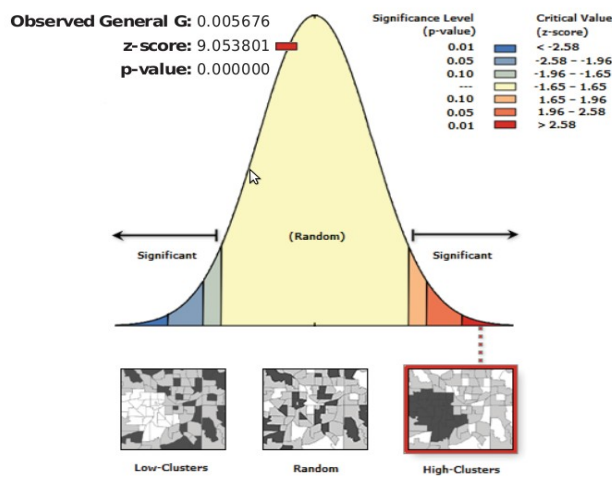
Para tomar en consideración la localidad de referencia se debe incluir de nuevo la localidad de referencia en el “vecindario”, como mencionado por Getis y Ord (1992). Una vez procesada esta lista, se elaboró el estadístico Z_i con la metodología anteriormente descrita con la función:

```
>LOCALG <- localG(x= “casos de cáncer agrupados 2012.2016”,
listw= n2listw(VECINOS, style= “B”, zero.policy=TRUE)
```

Esto establece un valor Z para cada localidad en la capa geográfica de INEGI. Para comprobar la significancia estadística global del proceso se utilizó una prueba global del estadístico.

Como se puede apreciar en el ejemplo de la Figura 1, el estadístico global G , muestra que tanto la dispersión o agrupamiento de “pockets” son estadísticamente significativos, los cuales se pueden discernir mediante patrones colorimétricos.

Figura 1. Significancia del estadístico global G



Given the z-score of 9.05, there is a less than 1% likelihood that this high-clustered pattern could be the result of random chance.

De acuerdo con Smith (2016), el valor de p menor al 0.01 indica que hay una probabilidad menor al 1% de que los agrupamientos formados por el código anterior y la información de las razones estandarizadas de mortalidad sea producto del azar.

6. Generación de los mapas de calor (Heat map en inglés)

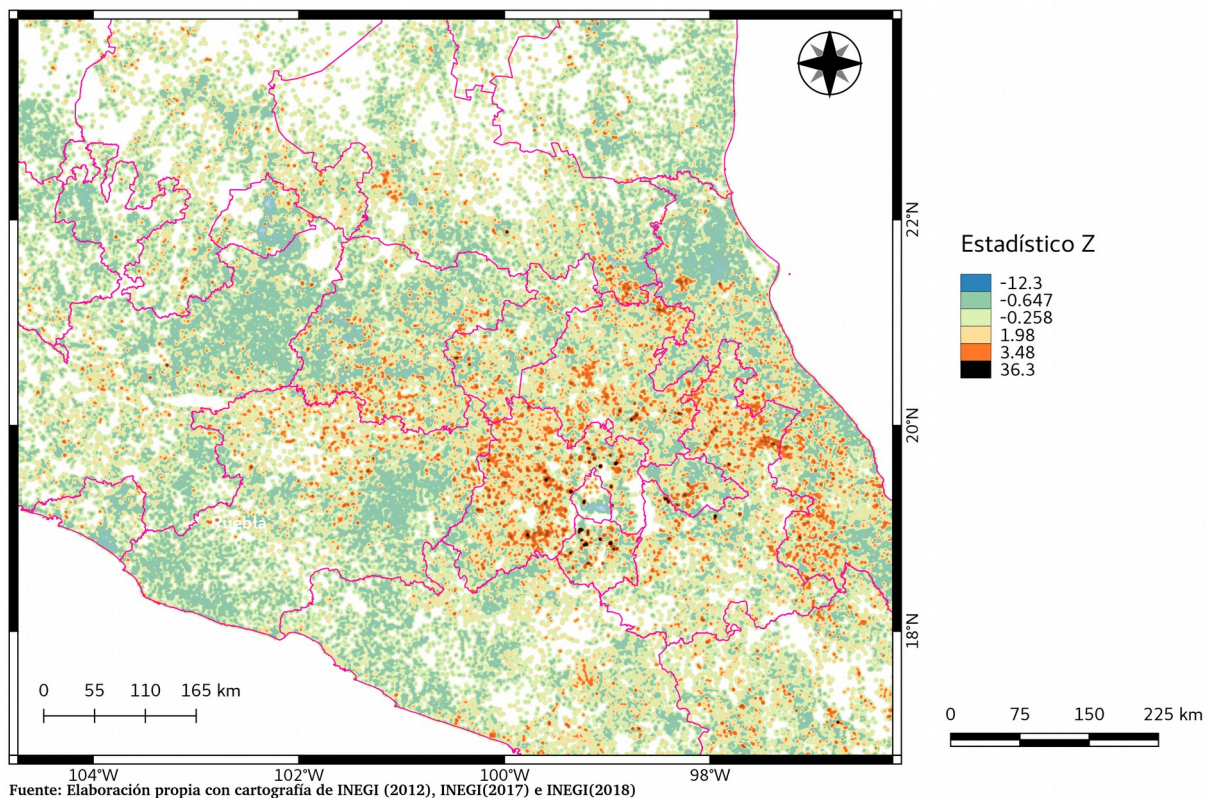
La información generada se anexó a una tabla con sus respectivas claves geográficas y se exportó a un archivo *.dbf para su uso en el software QGIS. La información fue anexada a la capa cartográfica “polígonos, localidades urbanas y rurales” y a los puntos geográficos de cada localidad obtenidos de las coordenadas del censo. Cabe destacar que la distribución de las razones estandarizadas de mortalidad y los valores z por localidad se comportan de una manera similar.

Con las coordenadas geográficas de las localidades disponibles en el censo de 2010 se generó un mapa de calor con los valores z para cada localidad. Aunque las estadísticas en las que se basa el mapa son robustas, hay que recordar que los radios de los kernel (el diámetro de cada punto en la imagen) en el mapa se asignan de manera arbitraria (para este análisis se asignó un valor de 2000

map units) y, por tanto, no reflejan la extensión real de la localidad en cuestión. Por otro lado, con el suavizamiento de los datos, los colores son más representativos de la mortalidad de cáncer para cada localidad. Sin embargo, hay que destacar que esto, a su vez, depende del gradiente de colores que se escoja para el raster (imagen georeferenciada). En este caso, se dividió a la población en percentiles y se seleccionaron los colores rojo para las localidades con REM ubicadas en el 5% superior (> Percentil 95%) y el color negro para el percentil 99%. Este mapa permite visualizar la información geográfica y realizar análisis estadísticos con mayor precisión.

Cabe destacar que la elección de los colores puede ser modificada acorde a las necesidades de cada proyecto. Por ejemplo, siguiendo esta metodología se desarrolló un mapa de la mortalidad por cáncer en el Eje Neovolcánico (Figura 2). Se pueden utilizar gradientes de colores, las localidades en amarillo y naranja representan las zonas arriba del percentil 98 para mortalidad por insuficiencia renal. En otras palabras, existen múltiples formas de expresar la información estadística en el gradiente de colores de un raster.

Figura 2. Mortalidad por cáncer en el Eje Neovolcánico, México.



7. Asociación del cáncer con la proximidad a la industria manufacturera

Con la información producida tanto con el método empírico de Bayes como con la de Getis y Ord (1992) es posible realizar diversos análisis espaciales para determinar si existe una asociación de ciertas enfermedades con otras variables georreferenciadas como, por ejemplo, la cercanía a la gran industria manufacturera, la minería, o la agroindustria, entre otras variables.

Para determinar una asociación estadística entre la mortalidad por cáncer y la cercanía a la industria, se calculó la distancia entre cada localidad y la industria manufacturera. Las coordenadas geográficas específicas para las principales industrias manufactureras en México, definidas como aquellas con más de 100 trabajadores, se obtuvieron del Directorio de Estadísticas Nacionales de Unidades Económicas (INEGI, 2018b).

Dado que la matriz resultante del comando spDists (Bivand et al, 2013) sería alrededor de 200 mil por más de siete mil, se optó por realizar el análisis de distancia dividiendo al país en las trece Regiones Hidrológicas Administrativas establecidas por la Comisión Nacional del Agua (CONAGUA, 2009). Después se unieron todas las matrices para determinar la relación de distancia de cada localidad con cada gran industria y calcular la distancia mínima de una localidad con una gran industria. Asimismo, se realizó la matriz con datos de la Comisión Nacional del Agua (CONAGUA 2018) que identifican los puntos de monitoreo más contaminados.

Debido a que la distribución de las razones estandarizadas de mortalidad con respecto a la distancia a un punto contaminado o a una gran industria no se comporta de manera normal, se evaluó la media para cada kilómetro con respecto a 7 kilómetros utilizando la prueba Mann-Whitney para determinar si la media por kilómetro era mayor entre más próxima se está a una gran industria (R Core Team, 2018).

8. Tipos de error en las fuentes de información

Debido a los problemas metodológicos para registrar con exactitud cada una de las defunciones a nivel nacional, podemos considerar que las causas de mortalidad registradas en estas 192,247 localidades en México es únicamente una muestra representativa de la mortalidad por cáncer en el país.

Para esta metodología, existen fuentes de error que no están directamente relacionados con la importación, combinación y análisis de las bases de datos del INEGI. El registro de datos erróneos en el acta de defunción es uno de los factores más importantes que afecta la confiabilidad de la información. Los errores más comunes en los certificados de defunción incluyen el registro erróneo de la causa subyacente de muerte, errores en la captura de la localidad de residencia; extravío del acta de defunción, y errores en la digitalización de las actas de defunción. Asimismo, los errores en los censos poblacionales, tanto por fallas metodológicas en las encuestas como por la falta de rigor en la captura de los datos también afectaría la confiabilidad de la información contenida en las bases de datos del INEGI.

Las imprecisiones en las coordenadas geográficas de las localidades del censo poblacional son aún otra fuente de error. Con nuestra metodología, los mapas generados pueden reducir a las ciudades con grandes extensiones a sólo un punto geográfico, y ciertas localidades con mayor proximidad a la gran industria o a otras localidades pueden aparentar estar más lejos de lo que realmente están. Una forma de evitar estos errores sería agregar los datos de mortalidad por manzana y utilizar metodologías de agregación (conocido como clustering en inglés) lo cual permitiría relacionar los factores ambientales y sociales con distintas afectaciones a la salud con mayor precisión.

9. Errores y límites de los ajustes y métodos

Es importante recalcar que los mapas que se producen con esta metodología no visualizan las razones estandarizadas de mortalidad, sino que estas son usadas en la metodología de Getis y Ord

(1992) como la variable x antes señalada. Esto implica que localidades con una razón estandarizada de mortalidad por debajo de lo esperado pueden tener valores Z altos si el “vecindario” en el que se encuentra está por arriba de la media global $G_i(d)$. El estadístico Z permite identificar aquellas localidades que se ubican en regiones donde se concentra la mortalidad de cáncer aunque la razón estandarizada de mortalidad de una localidad específica no sea superior a 1. El vecindario, en su conjunto, puede ser sistemáticamente superior a la media y así producir valores de Z altos para una localidad. Asimismo, es preciso aclarar que aunque el estadístico Z se utiliza para producir los mapas de calor, se utilizan las razones estandarizadas de mortalidad para calcular la relación entre la distancia a una gran industria y la mortalidad por cáncer. Aún no se incorporan los datos topográficos de elevación a esta metodología; ya que es un elemento importante dada la geografía de México, aún se está trabajando en un modo de incorporar la diferencia de elevación entre una gran industria y una localidad determinada.

Reconocimientos y agradecimientos: Este trabajo forma parte de un proyecto integral que realiza el *Colectivo Multidisciplinario por las Alternativas Locales (COMAL)* para desarrollar herramientas de información y articulación popular que sirvan para la defensa ambiental y social en México.

El presente trabajo no hubiera sido posible sin la asesoría de la Dra. Regina Montero y el Dr. Omar Arellano de la Universidad Nacional Autónoma de México. Asimismo, agradezco al Centro Fray Julián Garcés, al Mtro. Octavio Rosas Landa y a Adriana Martínez por apoyar en el desarrollo de esta metodología, y al Dr. Leopoldo Flores Romo del CINVESTAV-IPN y la Dra. Mussaret Zaidi Jacobson del Hospital General Dr. Agustín O’Horán por su revisión crítica del trabajo.

BIBLIOGRAFÍA

Bivand, R. (2017). "Creating Neighbours" en cran.r-project.org, 20 de noviembre de 2017. Última Consulta: [7 de Enero de 2018], Disponible en Internet: <https://cran.r-project.org/web/packages/spdep/vignettes/nb.pdf>

Bivand R, Gianfranco Piras (2015). Comparing Implementations of Estimation Methods for Spatial Econometrics. Journal of Statistical Software, 63(18), 1-36. Disponible en Internet: <http://www.jstatsoft.org/v63/i18/>.

Comisión Nacional del Agua [CONAGUA] (2009). "Regiones Hidrológicas Administrativas (Organismos de Cuencas)". [shape]. Última Consulta: [10 de abril de 2018]. Disponible en Internet: <http://www.conabio.gob.mx/informacion/gis/>

Comisión Nacional del Agua [CONAGUA] (2018). "Demana Química de Oxígeno" en Sistema Nacional de Información del Agua. [shape]. Última consulta: [15 de Mayo de 2018]. Disponible en Internet en: <http://sina.conagua.gob.mx/sina/>

Comisión Nacional de los Derechos Humanos (2017) Recomendación 10/2017. *Sobre la violación a los derechos humanos a un medio ambiente sano, saneamiento del agua y acceso a la información, en relación con la contaminación de los ríos Atoyac, Xochiac y sus afluentes; en agravio de quienes habitan y transitan en los municipios de San Martín Texmelucan y Huejotzingo, en el estado de Puebla; y en los municipios de Tepetitla, de Lardizabal, Nativitas e Ixtacuixtla de Mariano Matamoros, en el Estado de Tlaxcala*. Última consulta: [18 de Agosto de 2018], Disponible en Internet: http://www.cndh.org.mx/sites/all/doc/Recomendaciones/2017/Rec_2017_010.pdf

Getis, A., Ord, J. (1992). "The Analysis of Spatial Association by Use of Distance Statistics" en Geographical Analysis, vol. 24, num. 3. pp. 189-206. Última consulta: [6 de enero de 2018]. Disponible en Internet: <https://www.colorado.edu/geography/leyk/data/lit%205113%202016/Getis%20and%20Ord%201992.pdf>

Gómez-Rubio V.; J. Ferrándiz-Ferragud; A. López-Quílez (2005). Detecting clusters of disease with R. Journal of Geographical. Systems. 7, Number 2:189-206.

Hadley Wickham (2007). Reshaping Data with the reshape Package. Journal of Statistical Software, 21(12), 1-20. Disponible en Internet: <http://www.jstatsoft.org/v21/i12/>.

Hadley Wickham (2018). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.3.1. <https://CRAN.R-project.org/package=stringr>

Instituto Nacional de Geografía y Estadística [INEGI] (2012). "Censo de Población y Vivienda 2010: localidades". [datos: dbf]. Última consulta: [10 de abril de 2018]. Disponible en Internet: <http://www3.inegi.org.mx/rnm/index.php/catalog/71>

Instituto Nacional de Geografía y Estadística [INEGI] (2016). "Estadísticas de defunciones generales 2002-2015" [datos: dbf]. Última consulta: [10 de abril de 2018]. Disponible en Internet: <http://www.beta.inegi.org.mx/proyectos/registros/vitales/mortalidad/>

Instituto Nacional de Geografía y Estadística [INEGI] (2018). "Marco Geoestadístico Nacional". [shape]. Febrero 2018. Última Consulta: [10 de abril de 2018]. Disponible en Internet: http://www.inegi.org.mx/geo/contenidos/geoestadistica/m_geoestadistico.aspx

Instituto Nacional de Geografía y Estadística [INEGI] (2018b). "Directorio Estadístico Nacional de Unidades Económicas (Industrias Manufactureras)" [shape]. Última consulta: [10 de marzo de 2018]. Disponible en Internet en: <http://www.beta.inegi.org.mx/app/descarga/?ti=6#microdatos> (accessed march 10th 2018)

Instituto Nacional de Geografía y Estadística [INEGI] (2018c) "Red Hidrográfica escala 1:50000 edición 2.0". [shape]. Última consulta: [1 de julio de 2018]. Disponible en Internet: http://www.inegi.org.mx/geo/contenidos/reclat/hidrologia/regiones_hidrograficas.aspx

La Nueva Mañana (2017). "Se cumplen cinco años del juicio por agrotóxicos en barrio Ituzaingó Anexo". En *La Nueva Mañana*. Última Consulta: [14 de agosto de 2018]. Disponible en Internet en: <https://lmdiarrio.com.ar/noticia/19362/se-cumplieron-5-anos-del-juicio-por-agrotoxicos-en-barrio-ituzaingo-anexo>

Livingstone, G. (2016). "The villagers who fear herbicides" en *BBC News*. Última Consulta: [13 de agosto de 2018]. Disponible en Internet: <https://www.bbc.co.uk/news/health-36924361>.

Marx, K. (1975). *El Capital: crítica de la economía política*. Libro Primero, Volumen 1. México: Siglo XXI

Marx, K. (1982). *Escritos de Juventud*. México: FCE

Piñeros, M., Pardo, C., Gamboa, O., Hernández, G., (2010). Atlas de Mortalidad de Cáncer en Colombia. Última Consulta: [1 de febrero de 2018]. Diponible en Internet: https://www.researchgate.net/profile/Marion_Pineros/publication/311096266_Atlas_de_mortalidad_por_cancer_Colombia/links/583d751408ae8e63e614d75d/Atlas-de-mortalidad-por-cancer-Colombia.pdf

R Core Team (2018) . "R: A language and environment for statistical computing". *R Foundation for Statistical Computing*, Vienna. Disponible en Internat: <https://www.R-project.org/>

Smith, T.E., (2016). Notebook on Spatial Data Analysis [online]. Disponible en Internet: <http://www.seas.upenn.edu/~ese502/#notebook>. Última consulta: [7 de enero de 2018]

Washington State Department of Health [WSDH] (2012b). Guidelines for Working with Small Numbers. Última Consulta: [4 de febrero de 2018]. Disponible en Internet en: <https://www.doh.wa.gov/Portals/1/Documents/1500/SmallNumbers.pdf>

World Health Organization (2016). *International Statistical Classification of Diseases and related Health Problems, 10th revision (ICD-10)*. 5th edition. Geneva: WHO press.